# Simulations to Test Data Assimilation through Feedback Nudging

SANCHEZ BLAKE, Bethune-Cookman University

We carry out computer simulations to test the effectiveness of data assimilation by feedback nudging. This is done for the Lorenz equation and the Korteweg-de Vries (KdV) equation. The runs for the KdV equations are done using the Dedalus pseudo-spectral package in Jetstream, "an open-stack based cloud environment" [6].

**ACM Reference format:**

Sanchez Blake. 2017. Simulations to Test Data Assimilation through Feedback Nudging. 1, 1, Article 1 (August 2017), 11 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 INTRODUCTION

Many physical phenomena can be effectively depicted by scientific and numerical models of dynamical systems. These models can be used to anticipate the future behavior of the system, given that the initial conditions are known. When accurate initial conditions are not known coarse observations sustained over time can be injected into the model to determine "improved" estimates of the states and furthermore, give data about the "uncertainty" in the estimates. Dedalus is a software package for solving initial value problem (IVPs) of partial differential equations (PDEs) using the pseudo spectral method. The user specifies the system in a python script where the linear terms are parsed into a sparse matrix structure and the nonlinear terms are treated by fast Fourier



XXXX-XXXX/2017/8-ART1 $15.00

https://doi.org/10.1145/nnnnnnn.nnnnnnn

transforms. Dedalus is thus designed for domain problems on a regular physical space such as rectangles, boxes, or by change of variables. It is also set up to run in parallel with a simple MPI command.[3] IVPs are solved by spectrally discretizing the spatial space variables to convert the problem to a set of coupled ordinary differential equations (ODEs). Dedalus offers an extent of ODE integrators including multistep and Runge-Kutta. The same diagnostic regarding the solution as it is evolved can be specified in the same plain text manner as the system itself. Analyzing a geophysical fluid system, whose past, current, and future practices are of incredible interest, is administered by a group of partial differential equations (PDEs). The representing conditions are understood numerically to acquire future states beginning from an initial condition portraying the present condition of the system. Prior to the solution, this initial condition must be given, which, alongside the model conditions, controls the trajectory of the solution in space and time. Step by step instructions to get ready starting conditions with high caliber and exactness is getting an expanding consideration in the field of geophysical fluid dynamics.

The climate and the sea are two main fluids on Earth. Their activities influence the lives of people. To productively predict their future practices is vital. Since the inception of computers, higher-determination air models and maritime models have been produced. These models have demonstrated a surprising capacity to foresee some key wonders and to reenact some vital qualities of the atmosphere or ocean. The greater part of the models, be that as it may, require an entire and exact determination of the three-dimensional (3D) structure of the underlying condition of a considered framework. A part from conventional data, numerous new wellsprings of information, for example, satellite information, radar, profilers, and other remote-detecting gadgets, have turned out to be accessible. In any case, observations are not as numerous, and it is as yet difficult to gauge the greater part of the model's degrees of opportunity at a given time. Furthermore, the perceptions are unpredictably conveyed in space and time, and they have diverse structures of irregular mistake. In this way, an effective information absorption strategy is expected to join these unpredictable perceptions to produce the underlying conditions that are circulated on consistent model lattices. The improvement of data analysis system has predominantly experienced three phases: straightforward examination, measurable or ideal insertion, and variational investigation. Straightforward investigation strategies were for the most part utilized as a part of 1950$s$, when PCs were inaccessible or toward the starting stage. Basic investigation techniques were the most punctual bases of information osmosis. In the 1970$s$, factual contemplations were brought into the environmental information digestion. In light of these contemplations, a few types of ideal interjection were utilized to absorb perceptions into conjecture models. These ideal insertion investigation strategies were utilized as a part of numerous operational focuses around the world. In the 1990s, barometrical information digestion changed to variational strategies, specifically the three-and four-dimensional variational information absorption (3D-Var/4D-Var) by utilizing adjoint methods. The 3D-Var/4D-Var approaches endeavor to consolidate perceptions and foundation data in an ideal approach to create the most ideal gauge of the model beginning state. This method not just has wide applications for the absorption of environment and sea, additionally can be utilized for some different applications in numerical climate expectation.

Data assimilation plays a more vital part in numerical climate/weather forecast, and it is considered as an outskirts branch of air and maritime sciences.

Our approach is often called feedback nudging [5]. We consider the following pair of equations:

$$\frac{du}{dt} = F(u) \tag{1}$$

$$\frac{dU}{dt} = F(U) + \alpha I_h(U - u) \tag{2}$$

The solution to (1) from [2] represents the reality we wish to forecast. eg., the weather. Henceforth, we call it the reference solution. The problem is we do not have a complete initial condition for. Instead, we have continuous data over time for a projection $I_h K$ where $h$ represents the spatial resolution (notation adapted from 'Continuous data assimilation for the 2D Benard convection through velocity measurements alone' ).

## 2 METHODS

### 2.1 The Lorenz System

The Lorenz system is a toy model for the modular amplitudes in a nonlinear thermal convection problem. It was initially studied in the 1960s by Edward Lorenz, an American meteorologist. Lorenz noticed that even the smallest difference in the initial conditions of two solutions would result in them being completely uncorrelated

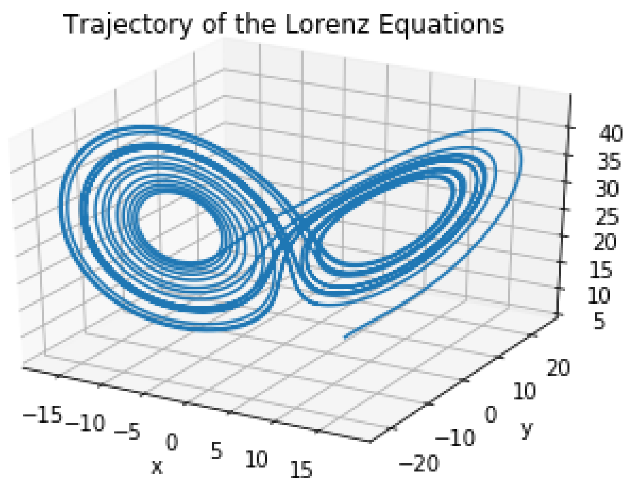


Fig. 1. This is a three dimensional plot of the trajectory of the Lorenz equation in the $xyz$ space. The sensitivity of solutions to perturbations of the initial data also has implications for numerical computations.

after a relatively short amount of time. This observation, sometimes referred to as the butterfly effect (it flaps its wings in Rio and dramatically alters the weather in New York), ushered in the modern theory of chaos in the field of dynamical systems. The phenomenon is now seen in most nonlinear systems. The Lorenz system is the following set of ordinary differential equations.

$$\begin{aligned} \frac{dx}{dt} &= -\sigma x + \sigma y \\ \frac{dy}{dt} &= rx - y - xz \\ \frac{dz}{dt} &= -bz + xy \end{aligned} \tag{3}$$

The role of $'u'$ will be played by $(x, y, z)$ while "$U$" will be played by $(X, Y, Z)$.

$$\begin{aligned} \frac{dX}{dt} &= -\sigma X + \sigma Y \\ \frac{dY}{dt} &= rX - Y - XZ - \alpha(Y - y) \\ \frac{dZ}{dt} &= -bZ + XY. \end{aligned} \tag{4}$$

The coarse projection of the reference solution is represented by just the y-component. We take the conventional trio of parameters, $\sigma = 10$, $r = 28$, and $b = 8/3$. The coarse projection of the reference solution is represented by just the $y$-component.

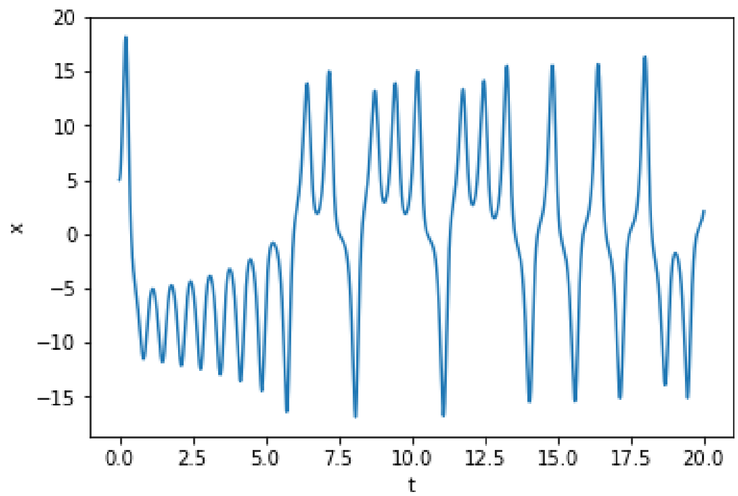


Fig. 2. A plot of $x$ vs $t$ for the same trajectory.

We now consider a dynamical system with spatial dependence, the Korteweg-de Vries (KdV) partial differential equation.

The KdV equation models surface waves in shallow water, and ion-acoustic waves in plasmas. In the former case u(x,t) has the meaning of the local elevation of the water surface. In addition to the surface shallow-water layers, KdV solutions modeling internal waves are also readily generated on interfaces between layers of different densities in stratified fluids.

$$u_t + u_{xxx} + uu_x + \gamma u = f \tag{5}$$

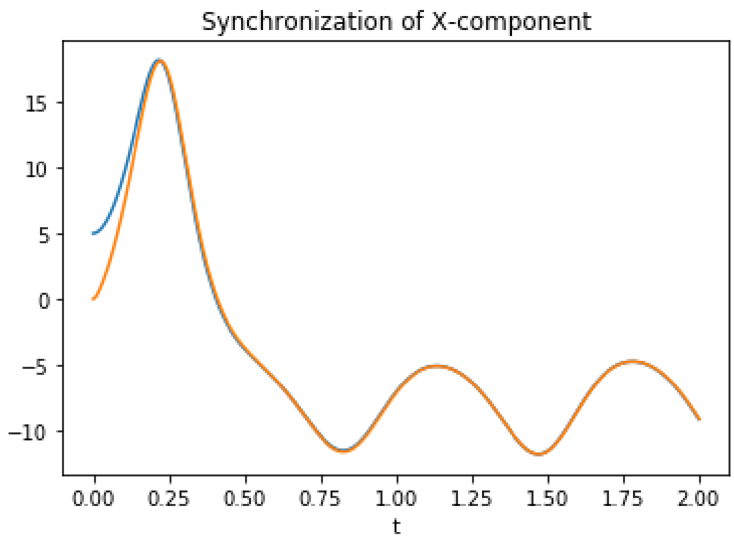


Fig. 3. The reference solution is in blue, X is in orange.

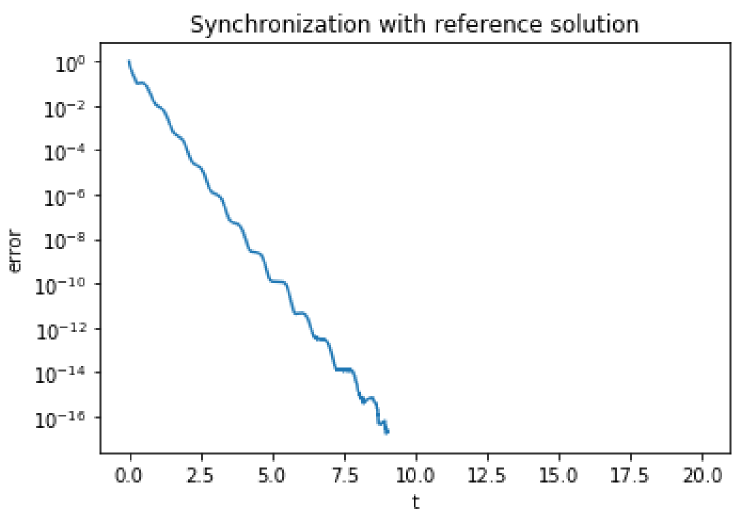


Fig. 4. error = $\frac{\sqrt{(x-X)^2+(y-Y)^2+(z-Z)^2}}{\sqrt{x^2+y^2+z^2}}$

We see that as $t$ increases the error is decaying at an exponential rate to 0.

$$U_t + U_{xxx} + UU_x + \gamma U = f - \alpha I_h(U - u) \tag{6}$$

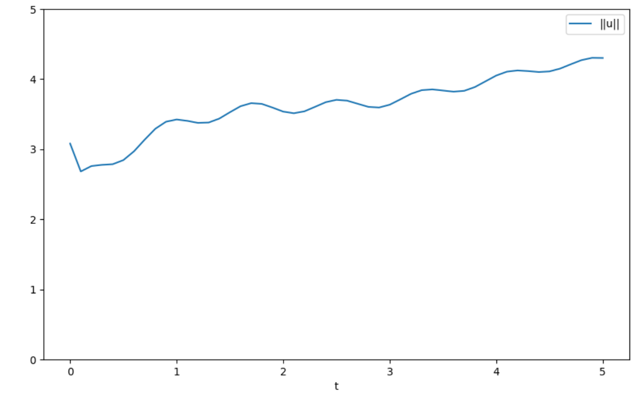


Fig. 5. The $L^2$-norm of the reference solution for $KdV$. Where $||u|| = \int_{-\pi}^{\pi} u^2 dx$.

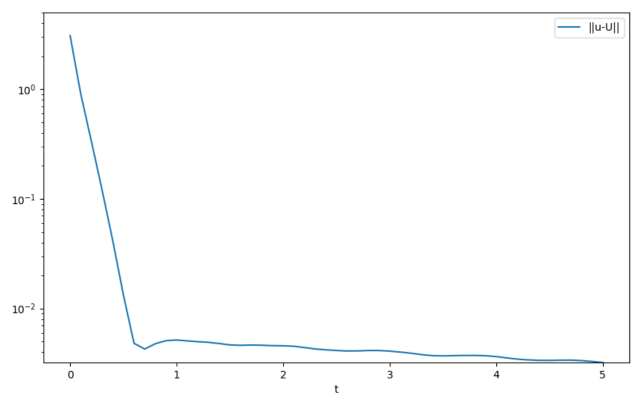


Fig. 6. The $L^2$-norm of $||u - U||$

## 2.2 Using the Fast Fourier Transform

Using separation of variables, we approximate our solution a truncated Fourier series. Taking the complex version, we write

$$S_M = \sum_{m=-M}^{M} U_m e^{imx}, \qquad U_{-m} = \bar{U}_m$$

We can approximate $u_x$ and $uu_x$ by differentiating term by term and find an approximation in uux by the two partial sums for the latter. Using a discrete convolution, we arrive at

$$uu_x \approx s_M s_M' = \sum_{m=-2M}^{2M} \Big( \sum_{\nu=-M}^{M} i\nu u_{m-\nu} u_\nu \Big) e^{imx}.$$

We can then write an ODE approximation of the KdV equation as

$$\frac{du_m}{dt} - im^3 u_m + \sum_{\nu=-M}^{M} i\nu u_{m-\nu} u_\nu + \gamma u_m = f_m$$

$$\frac{dU_m}{dt} - im^3 U_m + \sum_{\nu=-M}^{M} i\nu U_{m-\nu} U_\nu + \gamma U_m = \begin{cases} f_m + \alpha(U_m - u_m) & |m| < 8 \\ f_m & \text{else} \end{cases}$$

$$f = \sum_{m=-M}^{M} f_m e^{imx}, M = 128$$

The Fast-Fourier Transform, allows us to find the coefficients of $v = S_m S_M'$ much more efficiently than convolution We first take the coefficients of $u$ and $u_x$ and apply the inverse transform to get our grid values, $u(x_m)$ and $u_x(x_m)$. We can then multiply at each $(x_m)$ to get $u_x$ $(x_m)$, and then perform a forward transform to get our coefficients for the non-linear term. Each KdV eqn is then approximated by a system of ODEs (one for each coefficient). Each coefficient depends on time, so we get a time derivative. By chain rule, each derivative wrt $x$ produces a factor of

$im$, so $u_{xxx}$ has $-im^3$. The coefficient of the product is the convolution and we must include the damping term with gamma, and the force f on the right hand side (RHS). The same works for the forecasting equation, except for those modes with m < 8, we include the term with the data from the reference solution. That is only 8 out of 128 total modes used to synchronize the solutions, so the data is sparse.

## 3 CONCLUSION

"The goal of continuous data assimilation, and signal synchronization, is to use low spatial resolution observational measurements, obtained continuously in time, to accurately find the corresponding reference solution from which future predictions can be made. The motivating application of continuous data assimilation is weather prediction."[1] Using the Lorenz Equations, we recognized that our forecasting solution after a relatively short space of time and an arbitrary initial condition ended up being injected with our reference solution (with unknown initial condition).

## 4 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the Indiana University Jetstream team for providing me the opportunity to perform extensive research at this institution. I would also like to thank my immediate advisor Dr. M. Jolly, and by extension Dr. Christopher Connell of the Indiana University Department of Mathematics for facilitating and supervising the progress of my research and providing me a truly unforgettable experience that will help bolster my academic career.

## 5 APPENDIX

The Jetstream-cloud resource supports different modalities of utilization in aid of scientific research that doesn't exist elsewhere within the cyberinfrastructure sphere. These include: self-serve scholastic cloud administrations, based off virtual machine (VM) images given by the client or chosen from a library; constant (VMs) which bolster the delivery of science gateways. Jetstream additionally gives offices to distributing and sharing VM images by means of IU's determined computerized archive, IUScholarWorks, 'accessible via Digital Object Identifier. Access to interactive desktop sessions on Jetstream is available via VNC connections.' [4]. Along with the Jetstream team and Prof. Jolly (IU Mathematics Dept.), we are learning how to most effectively run mathematical software packages. The advantage of running in Jetstream, as opposed to doing so in batch systems, is the quick turn-around on preliminary runs in which certain parameters are being tested. The first software that I have tested is called Dedalus; it can solve a variety of partial differential equations modeling geophysical phenomena and is used in mathematics, applied mathematics, and engineering. The solutions depend on both time and a spatial variable, in some cases of three dimensions. One effort regarding turbulence requires a relatively long time average to see the patterns suggested by heuristic theories. In one on data assimilation, a projection of a solution from one system is inserted into another to synchronize a forecast. Dedalus, with its Python interface, enables users to specify the PDEs relatively easily, as opposed to developing a code from scratch. Moreover, it is a snap to run over multiple cores in Jetstream.

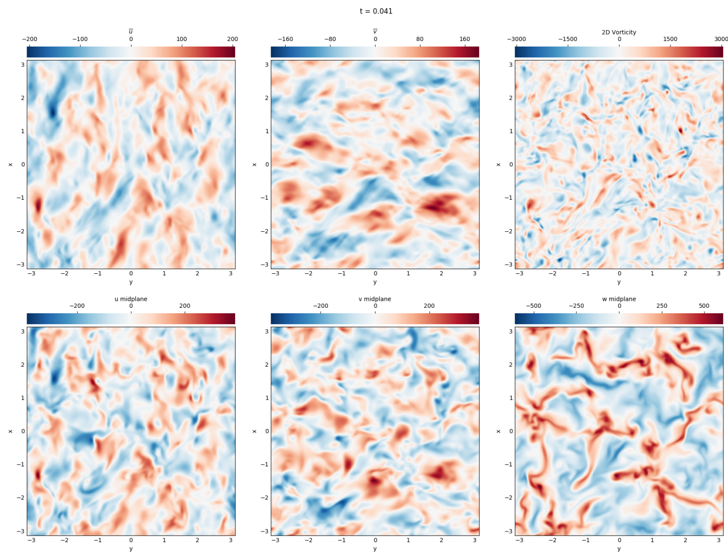


Fig. 7. Top is the vertically averaged velocity $\overline{u}$, $\overline{v}$ and the $2D$ vorticity for $3D$ Rayleigh Benard convection. Bottom is $u(x, y, 1/2)$ for all three components of $u$. One of the results of a team member using Dedalus on Jetstream.

## REFERENCES


[1] Abderrahim Azouani, Eric Olson, and Edriss S. Titi. Continuous data assimilation using general interpolant observables. *Journal of Nonlinear Science*, 2014.

[2] Aseel Farhat, Evelyn Lunasin, and Edriss S. Titi. Abridged continuous data assimilation for the 2D Navier-Stokes equations utilizing measurements of only one component of the velocity field. pages 1–27, 2015.

[3] J. S. Oishi D. Lecoanet B. P. Brown K. J. Burns, G. M. Vasil and E. Quataert. Dedalus: A Flexible Framework for Spectrally Solving Differential Equations.

[4] Richard Knepper, Eric Coulter, Marlon Pierce, Suresh Marru, and Sudhakar Pamidighantam. Using the Jetstream Research Cloud to provide Science Gateway resources. 2017.

[5] Evelyn Lunasin and Edriss S. Titi. Finite determining parameters feedback control for distributed nonlinear dissipative systems - a computational study. *arXiv:1506.03709*, (2000), 2015.

[6] Craig A Stewart, David Y. Hancock, Matthew Vaughn, Jeremy Fischer, Lee Liming, Nirav Merchant, Therese Miller, John Michael Lowe, Daniel Stanzione, Jaymes Taylor, and Edwin Skidmore. Jetstream - Performance, Early Experiences, and Early Results. In *Proceedings of the XSEDE16 Conference*, St. Louis, MO, 2016.